Review of Landwehr et al., 2021, ESD
Anonymous Referee #3

I find this paper uses an interesting approach that has a potentially high value and high impact for the ocean-atmosphere interdisciplinary research community. The paper takes the observations from the Antarctic Circumnavigation Expedition (ACE, austral summer 2016/2017) cruise and combines them with a sparse Principal Component Analysis (sPCA) to understand how different observed variables are linked together and to the general context (e.g. distance from land, cyclone activity, etc.). The paper is also very long, which makes reading and understanding the entire content of the paper and really getting into the new conclusions that result from this study extremely difficult.

I support this paper as a proof of concept for this approach, but I find the science questions posed (or hypotheses) and conclusions in the study are very weak. This paper should be published after the comments from the other reviewers and the comments below are addressed.

Most of the conclusions made using this very complex analysis are simplified statements of well known phenomena. So, I’m not sure what is the added value of this approach compared to what is already known. This is seen in the various “In summary” statements that come at the end of each section that focuses on the latent variables (LVs). This is seen most clearly in the summary for LV7 and LV10, which mostly put things into a seasonal and diurnal cycle context. I do not see what we have learned by using this “data science” approach. One way to address this would be to acknowledge in the abstract and very early in the study that there are no main scientific conclusions using data science in this study, but that this sets up the methodology that can be used in the future for this purpose.

The paper should be re-titled to more clearly reflect the paper content. The paper focuses on all of the aspects of the ACE cruise, not just biogeochemistry and physics. I would recommend something more general like “Understanding processes observed in the southern ocean-atmosphere system using ACE observations combined with data science”.

General comments:

Major comments:
I recommend that the authors work on shortening the paper by moving some of the very lengthy discussion into supplementary materials or into an annex to make this paper more readable. I would like the authors to get to the point of what was learned in addition to what is already known more quickly.

The authors should discuss how different timescales of processes that occur in nature that control the observed variables that were seen as a snapshot in space and time on the ship. Is it fair to group things into a data science approach variables that are observed in the atmosphere, ice, and ocean that have very different lifetimes and controlling factors that may not be co-located (i.e. relating them in the same space and time may give the wrong correlations/dependencies compared to what happens in nature)?

How do non-local processes get integrated into this approach? This is not currently clear for me.

The authors should expand their discussion of missing data and the influence this has on their analysis (as noted by reviewer 1).

**Minor comments:**

There are a few small typos as noted by reviewer 2. I suggest a careful re-reading before publication.